Construct to Associate: Cooperative Context Learning for Domain Adaptive Point Cloud Segmentation
Supplementary Materials

Guangrui Li
ReLER, AAII, University of Technology Sydney
guangrui.li@outlook.com

In Appendix A, we present more quantitative results for a better evaluation of the proposed method. Then, we complement more qualitative results in Appendix B for more intuitive comprehension and comparison. Finally, in Appendix C, we provide more implementation details.

A. More Quantitative Results

Table A. Experiments results of GTA-LiDAR [6] → SemKITTI [1] with SqueezeSegV2 as backbone.

Methods	Car	Pedestrian	mIoU
SqueezeSegV2 [6]	63.2	12.8	38.0
ePointDA [8]	70.7	12.9	41.8
SqueezeSegV2 + CCL (Ours)	**74.3**	**13.8**	**44.1**

Table B. Comparison with CosMix following their setup on SynLiDAR → SemanticKITTI.

Method	**car**	**bi.cle**	**mt.cle**	**truck**	**oth-v.**	**pers.**	**b.clst**	**m.clst**	**road**	**park.**	**sidew.**	**oth-g.**	**build.**	**fence**	**veget.**	**trunk**	**terra.**	**pole**	**traff.**	**mIoU**
ADDA [4]	52.5	4.5	11.9	0.3	3.9	9.4	27.9	0.5	52.8	4.9	27.4	0.0	61.0	17.0	57.4	34.5	42.9	23.2	4.5	23.0
Ent-Min [5]	58.3	5.1	14.3	0.3	1.8	14.3	**44.5**	0.5	50.4	4.3	34.8	0.0	48.3	19.7	67.5	34.8	**52.0**	33.0	6.1	25.8
ST [9]	62.0	5.0	12.4	1.3	9.2	16.7	44.2	0.4	53.0	2.5	28.4	0.0	57.1	18.7	69.8	**35.0**	48.7	32.5	6.9	26.5
PCT [7]	53.4	5.4	7.4	0.8	10.9	12.0	43.2	0.3	50.8	3.7	29.4	0.0	48.0	10.4	68.2	33.1	40.0	29.5	6.9	23.9
ST-PCT [7]	70.8	7.3	13.1	1.9	8.4	12.6	44.0	0.6	56.4	4.5	31.8	0.0	**66.7**	**23.7**	**73.3**	34.6	48.4	**39.4**	11.7	28.9
CosMix [3]	75.1	6.8	29.4	**27.1**	11.1	22.1	25.0	24.7	79.3	14.9	**46.7**	0.1	53.4	13.0	67.7	31.4	32.1	37.9	13.4	32.2
CosMix + CCL (Ours)	**77.2**	**8.1**	**33.3**	26.4	**14.2**	**23.3**	43.3	**25.6**	**83.1**	**16.1**	44.3	4.1	55.0	13.6	70.0	31.9	33.3	36.8	**16.1**	**34.5**

First, following [8], we conduct experiments on GTA-LiDAR → SemKITTI for a fair comparison with previous works. Inferred from Table A, we can observe that our method maintains its superiority against previous solutions, which is generally consistent with the gain achieved on the main benchmarks.

Analogously, in Table B, we present a more comprehensive comparison with CosMix [3], where integrate the proposed module into their proposed setup and framework. Apparently, our method outperforms it with a noticeable margin, *i.e.*, +2.3% mIoU, which further validates the effectiveness of the proposed approach.

B. More Qualitative Results.

In Fig. A, we provide more qualitative results on the validation set of SemKITTI [1] and compare them with a representative solution, *i.e.*, CoSMix, to allow a more intuitive comprehension and comparison.

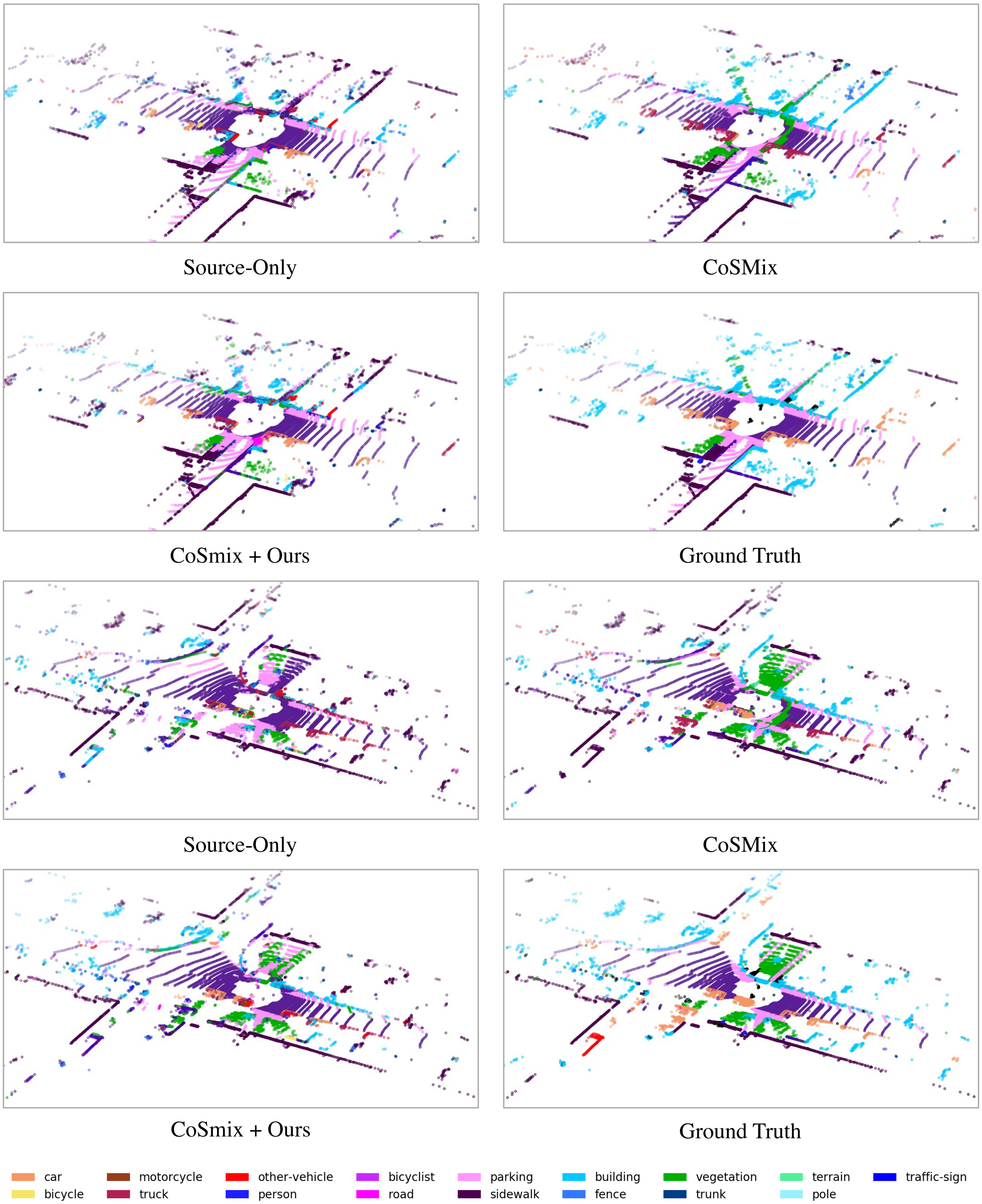

Figure A. More qualitative results on the validation set of SemKITTI [1].

C. More Implementation Details

Here we detail the implementation details for previous solutions.

CBST. It is trained for two stages, *i.e.*, the source-only training stage and the adaptation stage. The first stage trains on source samples only with the supervision loss. In the adaptation stage, the initial proportion for assigning pseudo labels is 10 %, and increases by 5 % every round, with 10 rounds in total. Each round consists of 5 epochs.

AdaptSeg. The weight for the adversarial training is 1e-3, and the objective for the domain adversarial training is the same as LSGAN [2] rather than the original form as we found its better stability for this adaptation task. The loss is imposed with feature embeddings of two domains, *i.e.*, output of the feature extractor and before the classifier.

PLCA. We adopt the original configuration according to their code, including the hyper-parameters.

MMD. We impose the MMD alignment loss on the feature space of two domains, with the weight of 0.01.

SqueezeSegV1. Following their paper and official code, we collect the frequency of noises for each spatial location in the 2D images projected with the spherical projection. Then we leverage the frequency map to guide the dropout on source samples as the input space, so as to mitigate the gap.

SqueezeSegV2. As SynLiDAR already provides the intensity channel, we adopt its domain alignment loss, *i.e.*, geodesic correlation alignment with the same weight.

LiDAR-NET. We adopt its official code and report the results on the two benchmarks.

CoSMix. As a data augmentation technique, we integrate it into the preprocessing procedures, which is performed before the spherical projection.

References

[1] J. Behley, M. Garbade, A. Milioto, J. Quenzel, S. Behnke, C. Stachniss, and J. Gall. SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences. In *ICCV*, 2019. 1, 2

[2] Xudong Mao, Qing Li, Haoran Xie, Raymond YK Lau, Zhen Wang, and Stephen Paul Smolley. Least squares generative adversarial networks. In *ICCV*, 2017. 3

[3] Cristiano Saltori, Fabio Galasso, Giuseppe Fiameni, Nicu Sebe, Elisa Ricci, and Fabio Poiesi. Cosmix: Compositional semantic mix for domain adaptation in 3d lidar segmentation. *ECCV*, 2022. 1

[4] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. In *CVPR*, 2017. 1

[5] Tuan-Hung Vu, Himalaya Jain, Maxime Bucher, Mathieu Cord, and Patrick Pérez. Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation. In *CVPR*, 2019. 1

[6] Bichen Wu, Xuanyu Zhou, Sicheng Zhao, Xiangyu Yue, and Kurt Keutzer. Squeezesegv2: Improved model structure and unsupervised domain adaptation for road-object segmentation from a lidar point cloud. In *ICRA*, 2019. 1

[7] Aoran Xiao, Jiaxing Huang, Dayan Guan, Fangneng Zhan, and Shijian Lu. Synlidar: Learning from synthetic lidar sequential point cloud for semantic segmentation. *AAAI*, 2022. 1

[8] Sicheng Zhao, Yezhen Wang, Bo Li, Bichen Wu, Yang Gao, Pengfei Xu, Trevor Darrell, and Kurt Keutzer. epointda: An end-to-end simulation-to-real domain adaptation framework for lidar point cloud segmentation. In *AAAI*, 2021. 1

[9] Yang Zou, Zhiding Yu, Xiaofeng Liu, B.V.K. Vijaya Kumar, and Jinsong Wang. Confidence regularized self-training. In *ICCV*, 2019. 1